\( |A \cup B| = |A| + |B| \) for disjoint \( A \) and \( B \)
\( |A \times B| = |A| \cdot |B| \)

**Strings:** A string over \( S \) is a sequence of elements over \( S \)

A \( k \)-string is a string of length \( k \)

substring, \( k \)-substring

permutation of \( S \)

\( k \)-permutation of \( S \) (\( k \)-substring of permutation of \( S \))

\( k \)-combination of \( S \) (\( k \)-subset of \( S \))

Number of permutations of \( S \): distinguish different orderings of identical elements of \( S \)

Number of combinations of \( S \): do not distinguish different orderings of identical elements of \( S \)

\( \binom{n}{k} = \frac{n!}{k!(n-k)!} \)

\( \binom{n}{k} = \binom{n}{n-k} \)

\( (x + y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} \)

\( 2^n = \sum_{k=0}^{n} \binom{n}{k} \)

\( \binom{n}{k} \geq \left( \frac{n}{k} \right)^k \)

\( \binom{n}{k} \leq \frac{n^k}{k^k} \leq \left( \frac{en}{k} \right)^k \) (Stirling's formula)

\( \binom{n}{k} \leq \frac{n^k}{k^n} \) (induction)

\( \left( \frac{n}{\lambda n} \right)^n \leq \left( \frac{1}{\lambda} \right)^{\lambda} = \left( \frac{1-\lambda}{\lambda} \right)^{1-\lambda} = 2^{nH(\lambda)} \) where \( H(\lambda) = -\lambda \log \lambda - (1 - \lambda) \log (1 - \lambda) \) and \( H(0) = 0 \)

A **probability distribution** \( \Pr \{ \} \) on a sample space \( S \) of elementary events is a map from **events** of \( S \) (subsets of \( S \)) to the real numbers such that

1. \( \Pr \{ A \} \geq 0 \) \( \forall A \)
2. \( \Pr \{ S \} = 1 \)
3. \( \Pr \{ A \cup B \} = \Pr \{ A \} + \Pr \{ B \} \) when \( A \cap B = \emptyset \)

**Implications:**

\( \Pr \{ \emptyset \} = 0 \)

complement \( \Pr \{ \bar{A} \} = \Pr \{ S - A \} \)

\( \Pr \{ A \cup B \} \leq \Pr \{ A \} + \Pr \{ B \} \)

For finite or countably infinite \( S \), \( \Pr \{ A \} = \sum_{s \in A} \Pr \{ s \} \)
Continuous random variable: $\Pr\{s\} = 1/|S|

Uniform continuous distribution: $\Pr\{c, d\} = \frac{d-c}{b-a}$ where $S = [a, b]$

Conditional probability of $A$ given $B$: $\Pr\{A\mid B\} = \frac{\Pr\{A\cap B\}}{\Pr\{B\}}$

A and $B$ are independent if $\Pr\{A \cap B\} = \Pr\{A\} \Pr\{B\}$
Equivalently, $\Pr\{A\mid B\} = \Pr\{A\}$

Bayes’s Law
Note $\Pr\{A \cap B\} = \Pr\{B\} \Pr\{A\mid B\} = \Pr\{A\} \Pr\{B\mid A\}$
So, $\Pr\{A\mid B\} = \frac{\Pr\{A\} \Pr\{B\mid A\}}{\Pr\{B\}}$
or $\Pr\{A\mid B\} = \frac{\Pr\{A\} \Pr\{B\mid A\}}{\Pr\{A\} \Pr\{B\mid A\} + \Pr\{\neg A\} \Pr\{B\mid \neg A\}}$

Discrete random variable:
A discrete random variable $X$ is a function from the sample space $S$ to the real numbers. $S$ must be discrete (or finite).
$\Pr\{X = x\} = \sum_{\{s \in S : X(s) = x\}} \Pr\{s\}$
$X = x$ means “the event $\{s \in S : X(s) = x\}$” (and similarly for < and >)
$\text{not}(X = x) \equiv X \neq x$ means “the event $\{s \in S : X(s) \neq x\} = S - (X = x)$”
Probability density: $f(x) = \Pr\{X = x\}$
Cumulative distribution: $g(x) = \Pr\{X \leq x\}$
$E[X] = \sum_x x \Pr\{X = x\}$

Alternative when $X$ is always a non-negative integer:
$E[X] = \sum_{x=0}^\infty x \Pr\{X = x\} = \sum_{x=1}^\infty x(\Pr\{X \geq x\} - \Pr\{X \geq x + 1\}) = \sum_{x=1}^\infty \Pr\{X \geq x\}$
$\text{Var}[X] = \text{E}[(X - E[X])^2] = E[X^2] - E^2[X]$

Linearity of Expectation:
For random variables, $X$ and $Y$, and real number $c$,
$E[\{X + Y\} = E[X] + E[Y]$
$E[cX] = cE[X]$

Joint probability density: $f(x, y) = \Pr\{X = x \text{ and } Y = y\}$
$X = x$ and $Y = y$ means “the event $\{s \in S : X(s) = x \text{ and } Y(s) = y\} = (X = x) \cap (Y = y)$”
$X = x$ or $Y = y$ means “the event $\{s \in S : X(s) = x \text{ or } Y(s) = y\} = (X = x) \cup (Y = y)$”
$E[XY] = \sum_x \sum_y xy \Pr\{X = x \text{ and } Y = y\}$

Two random variables $X$ and $Y$ are independent if and only if:
$\Pr\{X = x \text{ and } Y = y\} = \Pr\{X = x\} \Pr\{Y = y\}$ for all $x$ and $y$
If $X$ and $Y$ are independent, then $E[XY] = E[X]E[Y]$

Continuous random variable: $S \rightarrow \text{reals}$, $S$ continuous (or infinite)
$\Pr\{X \leq x\} = \Pr\{\{s : X(s) \leq x\}\}$ $\{s : X(s) = x\}$ is an event
(Hence, $\Pr\{X \leq x\} = \Pr\{\{s : X(s) \leq x\}\}$)
Probability density: more complicated, not just $\Pr\{X = x\}$
(For example, \( \Pr\{X = x\} = 0 \) always for the continuous uniform distribution)

Cumulative distribution: \( g(x) = \Pr\{X \leq x\} \)

\[
E[X] = \int_a^b x \frac{d\Pr\{X \leq x\}}{dx} \, dx \quad \text{(for } a \leq X \leq b) 
\]

\[
\text{Var}[X] = E[(X - E[X])^2] = E[X^2] - E^2[X]
\]

**Conditional Random Variable:**

If \( X \) is a random variable and \( A \) an event, then a **conditional random variable** \( X \mid A \) is another random variable defined by:

\[
(X \mid A)(s) = \frac{X(s)}{\Pr\{A\}} \forall s \in A \text{ and } (X \mid A)(s) = 0 \forall s \in S - A
\]

From this, and the definitions, one concludes that:

\[
(X \mid S) = X \\
\Pr\{(X \mid A) = x\} = \Pr\{(X = x) \mid A\} \\
\Pr\{(X \mid (A \cup B)) = x\} \Pr\{A \cup B\} = \Pr\{(X \mid A) = x\} \Pr\{A\} + \Pr\{(X \mid B) = x\} \Pr\{B\} \quad \text{for } A \text{ and } B \text{ disjoint events}
\]

The other definitions for random variables apply directly to conditional random variables.